Office of the State Controller

Enterprise Data Warehouse Feasibility
Agency Needs Assessment

Building Enterprise Access
For NC's Core Operation Needs

Table of Contents

Background

In September of 2005, the NC Office of the State Controller initiated a project with SAS to assess the needs of various state agencies for reporting and analysis of financial data. This project fell under the Data Warehousing Initiative of the Statewide Business Infrastructure Program (SBIP – now referred to as BEACON), and was one of four activities that addressed Data Warehousing.

The purpose of the project was to assess the statewide need for analysis of financial information. This was to be accomplished by interviewing business and technical representatives from various state agencies and departments to assess the following:

- What financial information do they need to carry out their respective missions?
- Are they able to get this information today?
- Do they have confidence in the quality of the information they get?
- What is the process for getting that information?
- What tools are used to obtain the information?
- Is that process effective and efficient?

By understanding the answers to these key questions and performing additional research and analysis, the study would evaluate whether there are potentially better solutions for getting statewide financial information to the people who need it to make effective business decisions.

The study was broken down into a series of phases and deliverables that encompassed the data collection, analysis and summarization portions of the study. This document serves as the final deliverable which summarizes the activities, key finding and recommendations for the study.

Activities

Deliverables

The key deliverables of the project were:

- Assessment Approach Document and Interview Questions
- Interview Followup and Documentation (First 7 Agencies)
- Interview Followup and Documentation (Second 5 Agencies)
- Needs Assessment Final Summary, Findings and Recommendations

The Assessment Approach and Document and Interview Questions were submitted and accepted September 15, 2005. Interviews were then conducted with the following six agencies through the remainder of September, October and early November, 2005:

- Department of Revenue
- Office of the State Controller
- Department of Commerce
- State Health Plan
- Office of State Management and Budget
- Office of State Personnel

After completing the above interviews and delivering the Interview Followup and Documentation for that group, a meeting was held between OSC and SAS where a mutual decision was made to alter the scope of the project and not conduct individual interviews for the second group of agencies. Instead, a conference was held at SAS on April 6, 2006 with representatives from 10 agencies which provided a more efficient and collaborative information exchange.

This document serves as the third and final deliverable for the project, and summarizes the activities, key findings and recommendations for the study.

Research and analysis

The period between the conclusion of the initial set of interviews ending in November, 2005 and the conference at SAS in April, 2006 was spent analyzing the findings from the interviews and performing additional research into statewide data warehousing needs.

In addition to the information obtained through direct interview sessions, SAS and OSC participated in a Data Warehousing study that was being conducted by the Technology Planning Group – a group of agency CIO's responsible for setting statewide technology vision and strategy. In addition, SAS was actively working with ITS on defining the technology infrastructure for providing a statewide data warehousing shared services utility offering, as well as completing the initial data warehousing prototype application for the OSP. All of these activities provided necessary input for assessing the statewide need and feasibility of data warehousing.

Key Findings

The following list summarizes the key findings from the Data Warehousing Financial Data Needs Assessment:

- Agencies have various reporting applications to access financial information
- Access to information is often cumbersome and not efficient
- The same data is being pulled from the same systems multiple times
- Data is commonly misrepresented or skewed due to lack of standardization
- Statewide data is often combined with agency or external data prior to analysis
- Reporting tools and expertise are not consistent across agencies

One key finding of the study was that agencies and departments are able to get most of the financial information they need to perform their mission. However, it was apparent throughout the study that each entity acted very autonomously in how they accomplished that. Each one seemed to have their own methodology and toolsets for accessing the raw data and reporting on it. This inconsistent approach leads to a number of issues.

First, since each entity is accessing the raw data directly (a process known as Extract/Transform/Load, or ETL), this can have a negative impact on the performance of the source system. In the case of financial data, the origin of much of that data is the statewide accounting system, NCAS. As more and more entities access that system to extract data for reporting and analysis, it can put an increasingly heavy workload on that system. Most operational systems like NCAS are designed for processing very large quantities of small sized transactions. The workloads associated with data extracts, are just the opposite – one transaction request for a very large amount of data.

A second issue with having multiple entities accessing the data using disparate tools and methodologies is that it can easily lead to inconsistent interpretations of the data. Since legacy systems have been changed and adapted over time to meet current needs, the way data is represented in those systems is often cryptic and difficult to understand. If the various encoding schemes and nuances are not all taken into account as the data is extracted and transformed, it can easily lead to decisions that are made based on incorrect information.

And finally, another issue with having multiple entities accessing the same source system in different ways (and especially relevant to the State of NC as it replaces those systems with SAP) is that whenever the sources systems are significantly modified or replaced, all of the ETL programs that access those systems also have to be modified or replaced. This can result in a very significant “flow down” effect for changes to the source system, and often results in unacceptable transition periods whereby information needed for decision making is not available.

The figure below provides an illustration of the State's current environment.

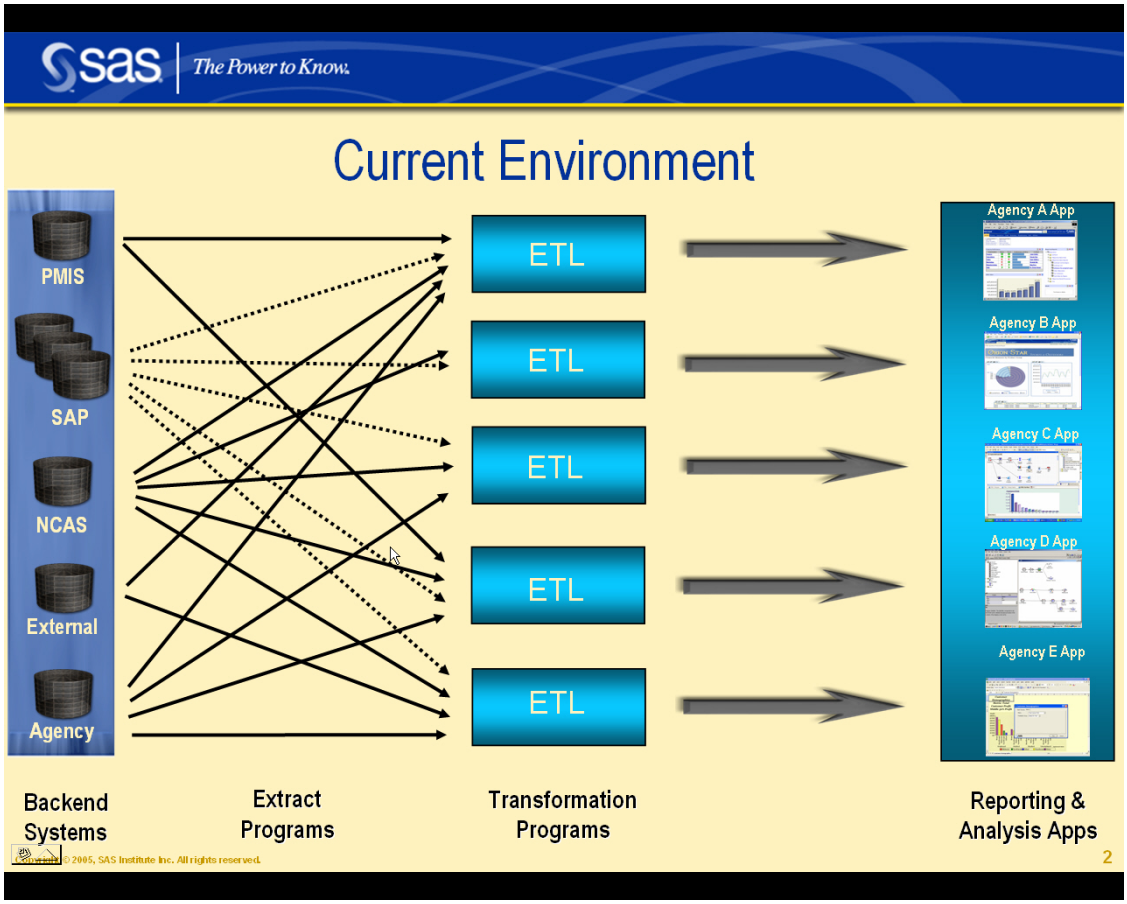

Figure 1: Financial Data Access – Current Environment

In addition to the technical challenges and data quality and consistency issues, the current process does not provide for economies of scale. A major focus of SB-991 is to provide greater efficiencies in the State's technology infrastructure. In the current environment, each entity acting autonomously greatly minimizes the opportunity for efficiency gains since each one is utilizing its own hardware infrastructure, software tools and methodologies. In addition there is almost no opportunity for sharing of expertise and components from one entity to another. The result is that the proverbial "wheel" is often re-invented many times over.

A final key finding is that each entity is often combining standard statewide financial data (such as NCAS data) with other agency specific or third party data in order to perform their analyses. This is important because it implies that the agencies do require some level of autonomy in order to perform the necessary analyses to support their mission.

Recommendations

In order to facilitate statewide access to financial information in a robust, consistent and cost effective manner, it is our recommendation that the State develop and adopt a "Statewide Information Strategy". This strategy will define and articulate, at a statewide level, how data assets are to be accessed, managed and controlled.

A major portion of the strategy should include the usage of standard tools and methodologies for statewide data access. The State has already determined that SAS is the statewide standard toolset for

Data Warehousing and Business Intelligence. To further maximize the economies of scale associated with using a standard toolset, the State should continue its efforts to build out the supporting architecture so that the SAS tools can be used on a statewide basis.

In addition, the State should develop an architectural roadmap for organizing and controlling statewide data assets. This roadmap should define the high level technical architecture, data architecture and governance model for statewide data assets. The roadmap should also define appropriate tools and methodologies that allow applicable consumers of the statewide data assets to access the data and combine it with other agency specific or third party data.

By providing a standardized toolset, architecture and methodology for managing statewide data assets, the State will be in a position to avoid much of the redundant, inconsistent and inefficient data access patterns that currently exist. Once the architecture is in place, it will be possible for the State to provide standard ETL processes from core statewide applications such as NCAS and SAP that will be efficient, consistent and easily maintained. If those ETL processes are centrally managed, then as the State transitions from NCAS to SAP it will be much easier to manage the downstream impact of the migration. More importantly, a statewide financial data warehouse could act as a crucial data integration hub that not only shields the end users of the information from constant changes to the source systems, but also provides a consolidated view of data from both systems during the often multi-year transition period.

The following diagram illustrates the proposed environment.

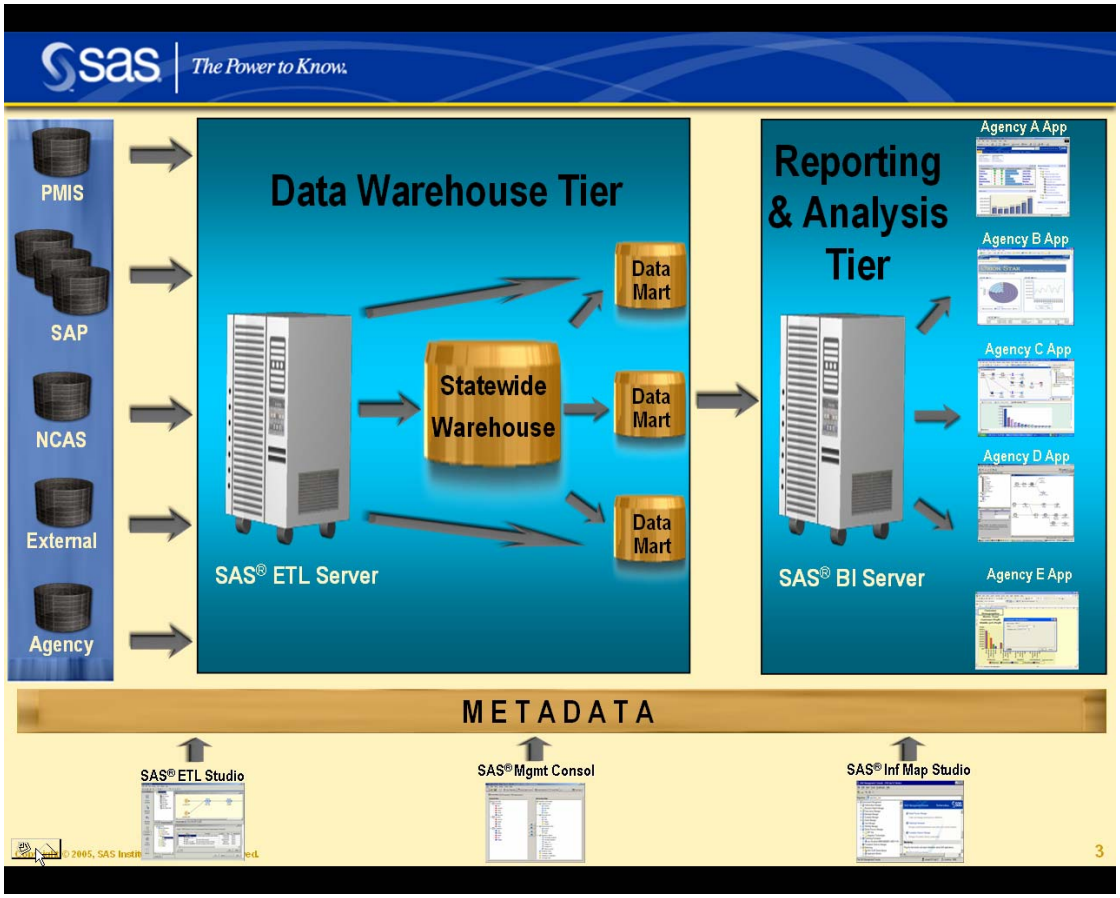

Figure 2: Financial Data Access – Proposed Environment